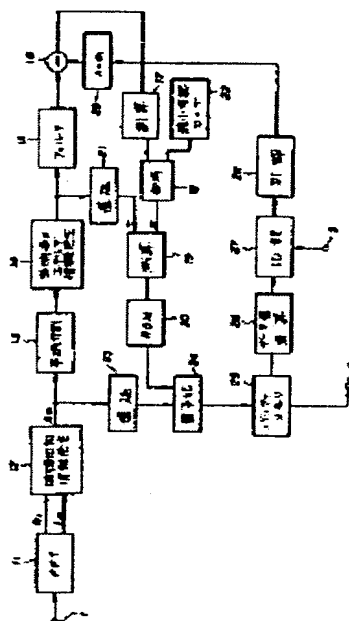


**DIGITAL SIGNAL ENCODING DEVICE****Publication number:** JP3117922**Publication date:** 1991-05-20**Inventor:** FUJIWARA YOSHIHITO; NISHIGUCHI MASAYUKI; UMEZAWA TOMOKO**Applicant:** SONY CORP**Classification:****- international:** H03M7/30; H03M7/30; (IPC1-7): H03M7/30**- european:****Application number:** JP19890255801 19890930**Priority number(s):** JP19890255801 19890930

Report a data error here

**Abstract of JP3117922**

**PURPOSE:** To minimize the degradation in tone quality regardless of a low bit rate by setting the allowable noise level for the same energy to a higher value in the case of a higher frequency by a noise level setting means. **CONSTITUTION:** A total sum detecting circuit 14 and a filter circuit 15 as the noise level setting means which sets the allowable noise level of each band unit based on the energy (or the peak value and the average value) of each of critical bands and a quantizing circuit 24 which quantizes components of each band with the number of bits assigned in accordance with the energy of each band and the level of the difference of the noise level setting means are provided. The allowable noise level for the same energy is set to a higher value in the case of the critical band having a higher frequency by the noise level setting means. Thus, the degradation in tone quality is minimized to reduce the bit rate.



Data supplied from the esp@cenet database - Worldwide

⑤Int.Cl.<sup>3</sup>

識別記号

庁内整理番号

⑬公開 平成3年(1991)5月20日

H 03 M 7/30

6832-5 J

審査請求 未請求 請求項の数 1 (全9頁)

⑭発明の名称 デジタル信号符号化装置

⑮特 願 平1-255801

⑯出 願 平1(1989)9月30日

⑰発明者 藤原 義仁 東京都品川区北品川6丁目7番35号 ソニー株式会社内  
 ⑰発明者 西口 正之 東京都品川区北品川6丁目7番35号 ソニー株式会社内  
 ⑰発明者 梅沢 知子 東京都品川区北品川6丁目7番35号 ソニー株式会社内  
 ⑱出願人 ソニー株式会社 東京都品川区北品川6丁目7番35号  
 ⑲代理人 弁理士 小池 晃 外2名

## 明細書

## 〔産業上の利用分野〕

## 1. 発明の名称

デジタル信号符号化装置

本発明は、入力デジタル信号の符号化を行う  
 デジタル信号符号化装置に関するものである。

## 2. 特許請求の範囲

入力デジタル信号を複数の周波数帯域に分割  
 すると共に、高い周波数帯域ほどバンド幅を広く  
 選定し、当該各バンド毎のエネルギーに基づいて各  
 バンド単位の許容ノイズレベルを設定するノイズ  
 レベル設定手段と、

上記各バンドのエネルギーと上記ノイズレベル設  
 定手段の差のレベルに応じたビット数で上記各バ  
 ンドの成分を量子化する量子化手段とを有するデ  
 ジタル信号符号化装置において、

上記ノイズレベル設定手段は高い周波数帯同一  
 のエネルギーに対する許容ノイズレベルを高く設  
 定するようにしたことを特徴とするデジタル信号  
 符号化装置。

## 〔発明の概要〕

本発明は、入力デジタル信号を複数の周波数  
 帯域に分割すると共に、高い周波数帯域ほどバン  
 ド幅を広く選定し、各バンド毎のエネルギーに基づ  
 いて各バンド単位の許容ノイズレベルを設定し、  
 各バンドのエネルギーと設定された許容ノイズレベ  
 ルの差のレベルに応じたビット数で各バンドの成  
 分を量子化するデジタル信号符号化装置において、  
 設定される許容ノイズレベルを高い周波数帯  
 同一のエネルギーに対して高く設定することにより、  
 音質劣化を最小限にしてビットレートを低減する  
 ことができるデジタル信号符号化装置を提供す  
 るものである。

## 〔従来の技術〕

オーディオ、音声等の信号の高効率符号化にお

## 3. 発明の詳細な説明

いては、オーディオ、音声等の入力信号を時間軸又は周波数軸で複数のチャンネルに分割すると共に、各チャンネル毎のビット数を適応的に割当てるビットアロケーション(ビット割当て)による符号化技術がある。例えば、オーディオ信号等の上記ビット割当てによる符号化技術には、時間軸上のオーディオ信号等を複数の周波数帯域に分割して符号化する帯域分割符号化(サブ・バンド・コーディング: SBC)や、時間軸の信号を周波数軸上の信号に変換(直交変換)して複数の周波数帯域に分割し各帯域毎で適応的に符号化するいわゆる適応変換符号化(ATC)、或いは、上記SBCといわゆる適応予測符号化(APC)とを組み合わせ、時間軸の信号を帯域分割して各帯域信号をベースバンド(低域)に変換した後複数の線形予測分析を行って予測符号化するいわゆる適応ビット割当て(APC-AB)等の符号化技術がある。

このような高能率符号化では、近年人間の聴覚上の特性におけるいわゆるマスキング特性を考慮

した高能率符号化の手法が盛んに試みられている。該マスキングの効果とは、ある信号によって他の信号がマスクされて聞こえなくなる現象をいうものであり、このマスキング効果には、時間軸上のオーディオ信号に対するマスキング効果と周波数軸上の信号に対するマスキング効果とがある。

ここでは周波数軸上のオーディオ信号に対するマスキング効果について述べる。例えば、ある周波数 $f$ の正弦波 $W$ があった場合、人間の聴覚によるマスキング効果を表現するマスキングスペクトル(マスキングカーブ)MSは、第7図のようになり、このマスキングスペクトルMSによって、図中斜線部で示す部分がマスキングされることになる。すなわち、該マスキングスペクトルMS内にノイズがあったとしても聞こえなくなるため、実際のオーディオ信号では、該マスキングスペクトルMS内のノイズは許容可能となる。このようなことから、上記正弦波 $W$ の場合の許容可能なノイズレベルは、第7図中jで示すレベル以下のレベルとなる。また、このとき、当該正弦波

$W$ の周波数 $f$ でマスキングの効果が最も高く、該正弦波 $W$ の周波数 $f$ から離れるにしたがってマスキングの効果が低くなる。

また、一般のオーディオ信号では、信号スペクトルSSは低域にエネルギーが集中していて、第8図のようになっていることが多い。この時のマスキングスペクトルMSは第8図中点線で示すようになる。このようなことから、当該信号スペクトルSSのオーディオ信号を再量子化(符号化)する際に発生するノイズのスペクトルが、第8図の斜線内に取まれば、該オーディオ信号のマスキングスペクトルMSによってノイズがマスクされ、人間の耳には音質の劣化は感じられないことになる。このように、オーディオ信号の量子化の際には、量子化ノイズのスペクトルを制御する手法が音質劣化に対して有効とされている。

(発明が解決しようとする課題)

しかし、上述のようなマスキング効果を利用した手法では、量子化のビット割当て数を決定する

際に、第8図で示したように信号スペクトルSSの形状に合わせてマスキングスペクトルMSの形状を決定しているので、一般に信号のエネルギーの低い高域部分にも多くのビット数を与えてしまい、結果としてビットレートを低くすることが困難となっている。

そこで、本発明は、上述のような実情に鑑みて提案されたものであり、オーディオ信号、音声信号等の信号スペクトル形状及びエネルギーを考慮した量子化ビット数で量子化を行うことにより、ビットレートを低くすることができるようにすると共に、低ビットレートであっても音質劣化を最小限に抑えることができるデジタル信号符号化装置を提供することを目的とするものである。

(課題を解決するための手段)

本発明のデジタル信号符号化装置は、上述の目的を達成するために提案されたものであり、例えば、第1図に示すように、入力デジタル信号を複数の周波数帯域に分割すると共に、高い周波

数帯域はとバンド幅を広く選定し、当該各バンド毎のエネルギーに基づいて各バンド単位の許容ノイズレベルを設定するノイズレベル設定手段としての総和検出回路14及びフィルタ回路15と、上記各バンドのエネルギーと上記ノイズレベル設定手段の差のレベルに応じたビット数で上記各バンドの成分を量子化する量子化回路24とを有するデジタル信号符号化装置において、例えば引算器16及び関数発生回路29を用いることで、上記ノイズレベル設定手段は高い周波数程同一のエネルギーに対する許容ノイズレベルを高く設定するようにしたものである。

#### 〔作用〕

本発明によれば、エネルギーの低いバンドすなわち高い周波数程許容ノイズレベルを大きくすることにより、高域での量子化の際の割当てビット数を少なくするようにしている。

#### 〔実施例〕

量子化する量子化回路24とを有するものであり、上記ノイズレベル設定手段は上記クリティカルバンドの高い周波数のバンド程同一のエネルギーに対する許容ノイズレベルを高く設定するようにしている。このため、第1図の装置では、後述するように、許容ノイズレベルを設定するための許容関数を、許容関数制御回路28に制御される許容関数発生回路29で発生させ、この許容関数に基づいて上記許容ノイズレベルを設定している。

その後、上記量子化回路24からの量子化出力は、バッファメモリ25を介して本実施例のデジタル信号符号化装置の出力端子2から出力されるようになる。

ここで、第1図に示す本実施例のデジタル信号符号化装置は、オーディオ信号等を高速フーリエ変換(FFT)して、時間軸の信号を周波数軸に変換した後、符号化(再量子化)を行ういわゆる適応変換符号化(ATC)を適用したものである。なお、本実施例装置は、特に、単位ブロック(フレーム)当たりのビットレートを一定化する

以下、本発明を適用した実施例について図面を参照しながら説明する。

本実施例のデジタル信号符号化装置は、オーディオ或いは音声等の入力デジタル信号を、例えば、帯域分割符号化(SBC)や、適応変換符号化(ATC)、適応ビット割当て(APC-AB)等により高効率符号化するものである。ここで、本実施例装置では、入力デジタル信号を複数の周波数帯域に分割すると共に、高い周波数帯域はとバンド幅を広く選定している。すなわち、後述する人間の聴覚特性を考慮したいわゆる臨界帯域幅(クリティカルバンド)で上記入力デジタル信号を帯域分割している。また、第1図に示すように、当該クリティカルバンドの各バンド毎のエネルギー(又はピーク値、平均値)に基づいて各バンド単位の許容ノイズレベルを設定するノイズレベル設定手段としての総和検出回路14及びフィルタ回路15と、上記各バンドのエネルギーと上記ノイズレベル設定手段の差のレベルに応じて割当てられたビット数で上記各バンドの成分を量

ようなビットレート調整(いわゆるビットパッキング)を行う一具体例を示している。

すなわち、第1図において、入力端子1には、例えばオーディオ信号が供給されており、この時間軸上のオーディオ信号が高速フーリエ変換回路11に伝送される。この高速フーリエ変換回路11では、上記時間軸上のオーディオ信号が所定時間(フレーム)毎に周波数軸上の信号に変換され、実数成分値 $R_e$ と虚数成分値 $I_m$ とからなるFFT係数が得られる。これらFFT係数は振幅位相情報発生回路12に伝送され、当該振幅位相情報発生回路12では上記実数成分値 $R_e$ と虚数成分値 $I_m$ とから振幅値 $A_m$ と位相値とが得られて、この振幅値 $A_m$ の情報が出力されるようになる。すなわち、一般に人間の聴覚は周波数領域の振幅(パワー)には敏感であるが、位相についてはかなり鈍感であるため、本実施例では上記振幅位相情報発生回路12の出力から上記振幅値 $A_m$ のみを取り出し、これを本発明実施例での入力デジタル信号としている。

このようにして得られた振幅値  $A_m$  等の入力デジタル信号は、帯域分割回路 13 に伝送される。当該帯域分割回路 13 では、上記振幅値  $A_m$  で表現された入力デジタル信号をいわゆる臨界帯域幅（クリティカルバンド）に分割している。当該クリティカルバンドとは、人間の聴覚特性（周波数分析能力）を考慮したものであり、例えば 0 ~ 16 kHz を 24 バンドに分け、高い周波数帯域ほどバンド幅を広く選定しているものである。すなわち、人間の聴覚は、一種のバンドパスフィルタのような特性を有していて、この各フィルタによって分けられたバンドを臨界帯域と呼んでいる。ここで、第 2 図に上記クリティカルバンドを示す。ただし、当該第 2 図では図示を簡略化するため、上記クリティカルバンドのバンド数を 12 バンド（ $B_1 \sim B_{12}$ ）で表現している。

上記帯域分割回路 13 でクリティカルバンドに分割された各バンド（例えば 24 バンド）毎の上記振幅値  $A_m$  は、それぞれ上記総和検出回路 14 に伝送される。当該総和検出回路 14 では、上記

各バンド毎のエネルギー（各バンドでのスペクトル強度）が、各バンド内のそれぞれの振幅値  $A_m$  の総和（振幅値  $A_m$  のピーク又は平均或いはエネルギー総和）をとることにより求められる。該総和検出回路 14 の出力すなわち各バンドの総和のスペクトルは、一般にパークスペクトルと呼ばれ、この各バンドのパークスペクトル  $S_B$  は例えば第 3 図に示すようになる。

ここで、上記パークスペクトル  $S_B$  のマスキングに於ける影響を考慮するため、上記パークスペクトル  $S_B$  に所定の重みづけの関数を畳込む（コンボリューション）。このため、上記総和検出回路 14 の出力すなわち上記パークスペクトル  $S_B$  の各値は、フィルタ回路 15 に送られる。該フィルタ回路 15 は、第 4 図に示すように、入力端子 100 からの入力データを順次遅延させる遅延素子（ $z^{-1}$ ）101<sub>1</sub>, 101<sub>2</sub>, ..., 101<sub>n-1</sub> ~ 101<sub>n</sub>, ..., 101<sub>23</sub>, 101<sub>24</sub>（クリティカルバンドに対応した例えば 24 個の遅延素子）と、各遅延素子 101<sub>1</sub> ~ 101<sub>24</sub> からの出力にフィ

ルタ係数（重みづけの関数）を乗算する例えば 24 個の乗算器 102<sub>1</sub>, 102<sub>2</sub>, ..., 102<sub>n-1</sub> ~ 102<sub>n</sub>, ..., 102<sub>23</sub>, 102<sub>24</sub> と、総和加算器 104 とから構成されるものである。この時、乗算器 102<sub>1</sub> ~ 102<sub>n</sub> において、例えば、乗算器 102<sub>1</sub> でフィルタ係数 0.0000086 を、乗算器 102<sub>2</sub> でフィルタ係数 0.0019 を、乗算器 102<sub>3</sub> でフィルタ係数 0.15 を、乗算器 102<sub>4</sub> でフィルタ係数 1 を、乗算器 102<sub>5</sub> でフィルタ係数 0.4 を、更に乗算器 102<sub>6</sub> でフィルタ係数 0.06 を、乗算器 102<sub>7</sub> でフィルタ係数 0.007 を各遅延素子の出力に乗算することにより、上記パークスペクトル  $S_B$  の畳込み処理が行われる。該畳込み処理により、第 3 図中点線で示す部分の総和がとられる。

ところで、上記パークスペクトル  $S_B$  のマスキングスペクトル（許容可能なノイズスペクトル）を算出する場合の後述する許容可能なノイズレベルに対応するレベル  $\alpha$  においては、このレベル  $\alpha$  が小さいとマスキングスペクトル（マスキングカ

ーブ）が下降することになり、結果として量子化回路 24 の量子化の際に割り当てるビット数を増やさなければならないようになる。逆に、上記レベル  $\alpha$  が大きいとマスキングスペクトルが上昇することになり、結果として量子化の際の割り当てるビット数を減少することができるようになる。なお、上記許容可能なノイズレベルに対応するレベル  $\alpha$  とは、後述するように、逆コンボリューション処理を行うことによってクリティカルバンドの各バンド毎の許容ノイズレベルとなるようなレベルである。また、一般にオーディオ信号等では、高域部分のスペクトル強度（エネルギー）が小さい。したがって本実施例においては、これらのことを考慮して、エネルギーの小さい高域にいく程、上記レベル  $\alpha$  を大きくし、当該高域部分のビット割当て数を減らすようにしている。このようなことから、上記ノイズレベル設定手段は高い周波数程同一のエネルギーに対する上記レベル  $\alpha$  を高く設定している。

すなわち、本実施例装置では、上記許容ノイズ

レベルに対応するレベル $\alpha$ を算出し、該レベル $\alpha$ が高域程高くなるように制御している。このため、上記フィルタ回路15の出力は引算器16に送られる。該引算器16は、上記畳込んだ領域でのレベル $\alpha$ を求めるものである。ここで、上記引算器16には、上記レベル $\alpha$ を求めるための許容関数（マスキングレベルを表現する関数）が供給される。該許容関数を増減させることで上記レベル $\alpha$ の制御を行っている。該許容関数は、後述する関数制御回路28によって制御された関数発生回路29から供給されている。

すなわち、許容ノイズレベルに対応するレベル $\alpha$ は、クリティカルバンドのバンドの低域から順に与えられる番号を $i$ とすると、第(1)式で求めることができる。

$$\alpha = S \cdot (n - ai) \dots\dots (1)$$

この第(1)式において、 $n$ 、 $a$ は定数で $a > 0$ 、 $S$ は畳込み処理後のバークスペクトルの強度であ

当該減算器19には、上記総和検出回路14の出力すなわち前述の総和検出回路14からのバークスペクトル $SB$ が、遅延回路21を介して供給されている。したがって、この減算器19で上記マスキングスペクトルとバークスペクトル $SB$ との減算演算が行われることで、第5図に示すように、上記バークスペクトル $SB$ は、該マスキングスペクトル $MS$ の各レベルで示すレベル以下がマスキングされることになる。

当該減算器19の出力は、ROM20を介して量子化回路24に供給されている。上記量子化回路24では、この減算器19の出力に応じた割合でビット数で、遅延回路23を介して供給されている振幅値 $A_m$ の量子化を行っている。すなわち、換言すれば、当該量子化回路24では、上記クリティカルバンドの各バンドのエネルギーと上記ノイズレベル設定手段の差のレベルに応じて割合で与えられたビット数で上記各バンドの成分を量子化することになる。なお、上記遅延回路21は上記合成回路18以前の各回路での遅延量を考慮して上記

り、第(1)式中 $(n - ai)$ が許容関数となる。ここで、上述した様に、エネルギーの少ない高域からビット数を減らす方が全体のビット数削減に有利であるため、本実施例では $n = 38$ 、 $a = 1$ としており、この時の音質劣化はなく、良好な符号化が行えた。

上述のようにして、上記レベル $\alpha$ が求められ、このデータは、割算器17に伝送される。当該割算器17では、上記畳込み処理された領域でのレベル $\alpha$ を逆コンボリューションするためのものである。したがって、この逆コンボリューション処理を行うことにより、上記レベル $\alpha$ から、マスキングスペクトルが得られるようになる。すなわち、このマスキングスペクトルが許容ノイズスペクトルとなる。なお、上記逆コンボリューション処理は、複雑な演算を必要とするが、本実施例では簡略化した割算器17を用いて逆コンボリューションを行っている。

次に、上記マスキングスペクトルは、合成回路18を介して減算器19に伝送される。ここで、

総和検出回路14からのバークスペクトル $SB$ を遅延させ、上記遅延回路23は上記ROM20以前の各回路での遅延量を考慮して上記振幅値 $A_m$ を遅延させるために設けられている。また、上記ROM20は量子化の際の所定時間毎の上記減算器19の出力を一時格納して送り出すために設けられている。

ここで、上述した合成回路18での合成の際には、最小可聴カーブ発生回路22から供給される第6図に示すような人間の聴覚特性であるいわゆる最小可聴カーブ（等ラウドネス曲線） $RC$ を示すデータと、上記マスキングスペクトル $MS$ とを合成することができる。したがって、この最小可聴カーブ $RC$ とマスキングスペクトル $MS$ とを共に合成することで、許容ノイズレベルはこの図中斜線で示す部分までとすることができ、量子化の際に図中斜線で示す部分の割合でビット数を減らすことができるようになる。なお、この第6図は、前述の第2図に示したクリティカルバンドで表されており、信号スペクトル $SS$ も同時に示してい

る。

また、上記バッファメモリ25からのデータは、データ量演算回路26によってデータ量が求められた後、比較回路27に送られる。当該比較回路27では、上記データ量と端子3からのビットレート調整のための1フレーム内ビット数の目標値のデータとが比較され、その比較結果が上述した関数制御回路28に伝送されている。したがって、当該関数制御回路28は、上記関数発生回路28を制御することにより、該関数発生回路28から、上記レベル $\alpha$ を求めるための許容関数を発生させるようにすると共に、ビットレート調整のための関数をも発生させるようにしている。

上述したように、本実施例のデジタル信号符号化装置においては、エネルギーの小さい高域にいく程、許容ノイズレベルを大きくし、当該高域部分のビット割当て数を減らすようにしているため、量子化の際の割当てビット数を減らすことができるようになる。

なお、本実施例においては、上述したビットレ

ート調整処理及び／又は最小可聴カーブの合成処理を行わない構成とすることもできる。すなわち、上記ビットレート調整を行わない構成の場合には、データ量演算回路26、比較回路27、関数制御回路28が不要となり、関数発生回路29からの許容関数は、固定され（例えば許容関数が38-1のように固定される）ることになる。また、最小可聴カーブを合成しない構成の場合には、最小可聴カーブ発生回路22、合成回路18が不要となり、上記引算器16からの出力は、割算器17で逆コンボリューションされた後、すぐに減算器19に伝送されることになる。

本発明は、上述した第1図の実施例装置のように、入力デジタル信号を高速フーリエ変換して処理する適応変換符号化の他に、例えば、帯域分割符号化等を行う装置にも適用することができる。この場合には、信号をバンドパスフィルタ等で帯域分割して、上記ノイズレベル設定手段は高い周波数の帯域程同一のエネルギーに対する許容ノイズレベルを高く設定するようにするものとなる。当

該帯域分割符号化等の場合も上述同様の効果を得ることができるようになる。

#### 〔発明の効果〕

本発明のデジタル信号符号化装置においては、オーディオ信号、音声信号等の信号スペクトル形状及びエネルギーを考慮し、設定される許容ノイズレベルを高い周波数程同一のエネルギーに対して高く設定することにより、高い周波数でマスキングスペクトルを高くすることができるようになり、量子化の際の割当てビット数を少なくする（ビットレートを低減する）ことができるようになる。したがって、少ないビット数であっても音質劣化を最小限に抑えることができることになる。

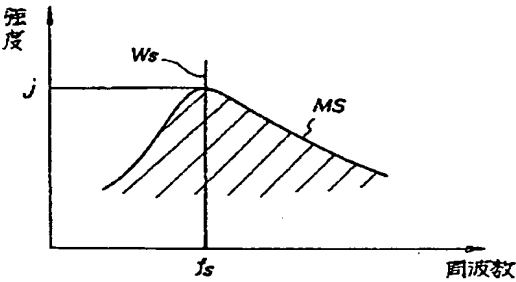
#### 4. 図面の簡単な説明

第1図は本発明の一実施例のデジタル信号符号化装置の概略構成を示すブロック回路図、第2図はクリティカルバンドを示す図、第3図はバークスペクトルを示す図、第4図はフィルタ回路を

示す回路図、第5図はマスキングスペクトルを示す図、第6図は最小可聴カーブ、マスキングスペクトルを合成した図、第7図は人間の聴覚のマスキングスペクトルを示す特性図、第8図はオーディオ信号のスペクトルを示す特性図である。

- 11.....高速フーリエ変換回路
- 12.....振幅位相情報発生回路
- 13.....帯域分割回路
- 14.....総和検出回路
- 15.....フィルタ回路
- 16.....引算器
- 17.....割算器
- 18.....合成回路
- 19.....減算器
- 20.....ROM
- 21, 23...遅延回路
- 22.....最小可聴カーブ発生回路
- 24.....量子化回路
- 25.....バッファメモリ

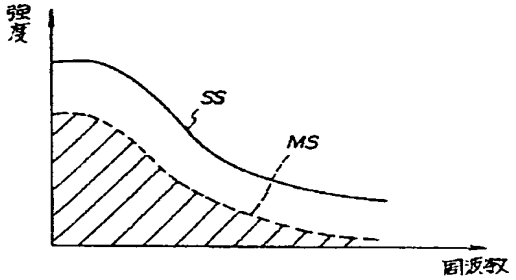
- 26.....データ量演算回路
- 27.....比較回路
- 28.....閾値制御回路
- 29.....閾値発生回路



人間の聴覚のマスクングスペクトル

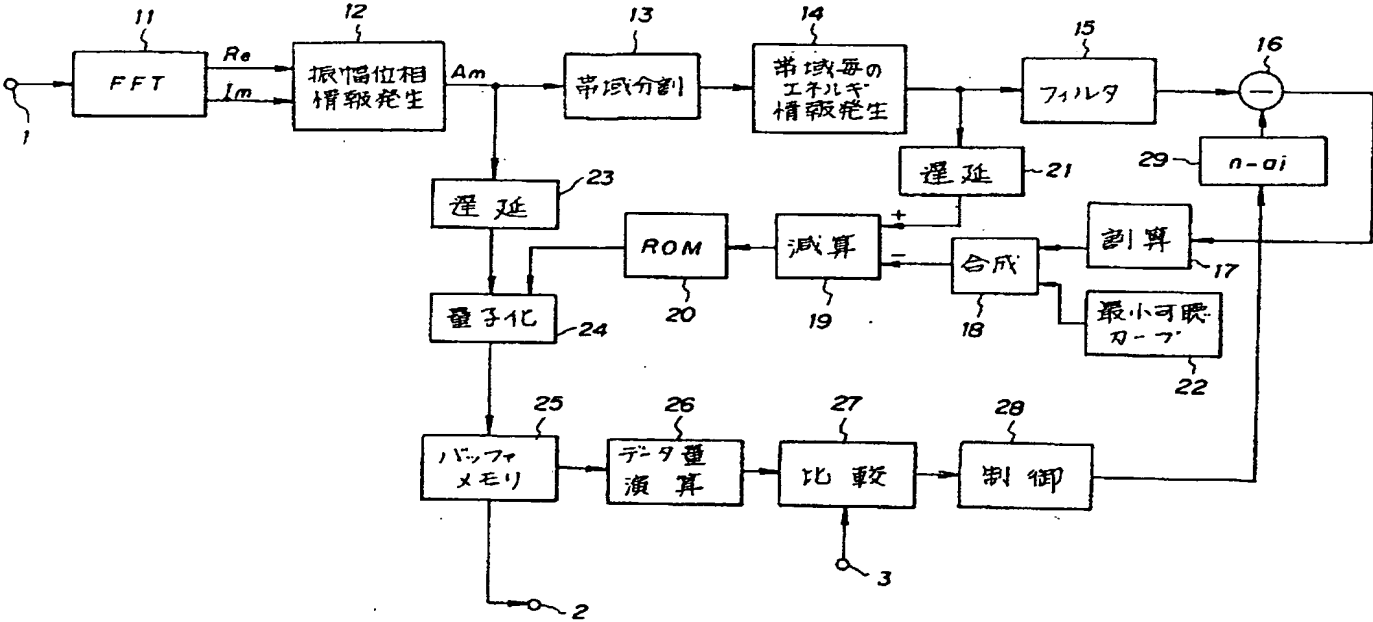
第 7 図

特許出願人      ソニー株式会社  
代理人 弁理士    小池 晃  
                    同      田村 榮一  
                    同      佐藤 勝



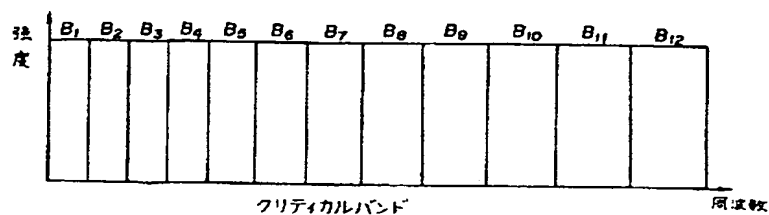
オーディオ信号のスペクトル

第 8 図

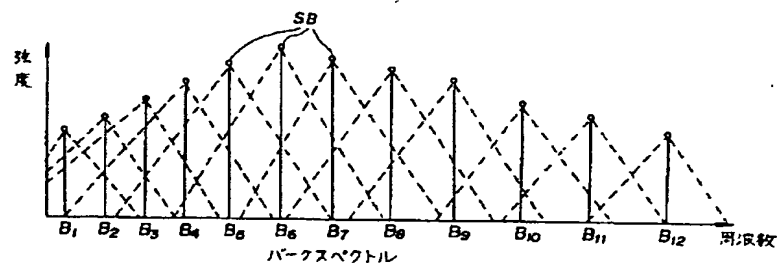


実施例装置  
第 1 図

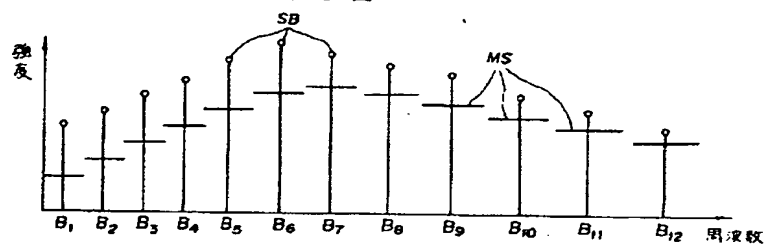




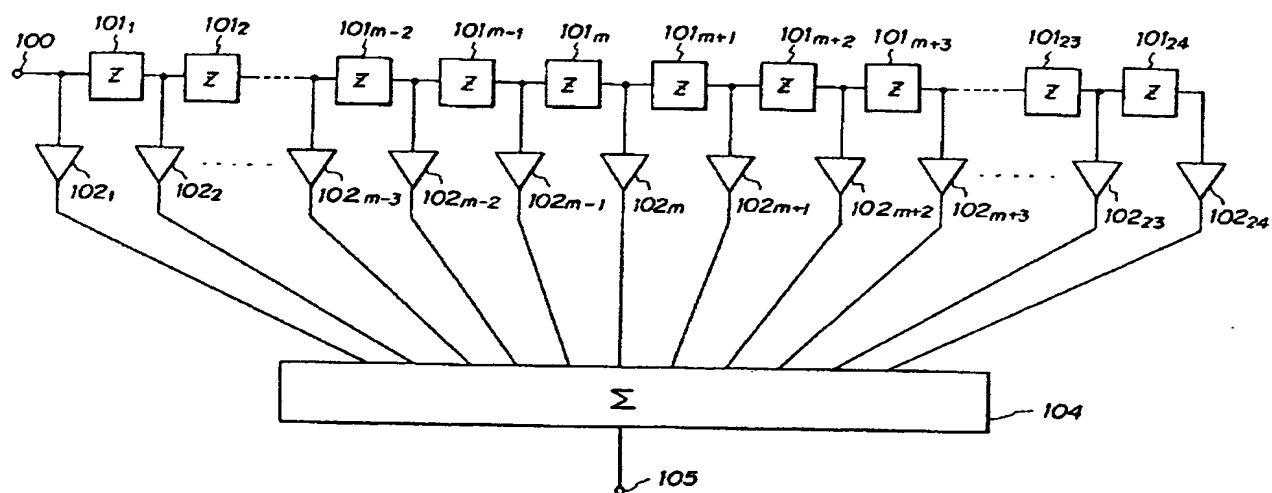
第 2 図



第 3 図

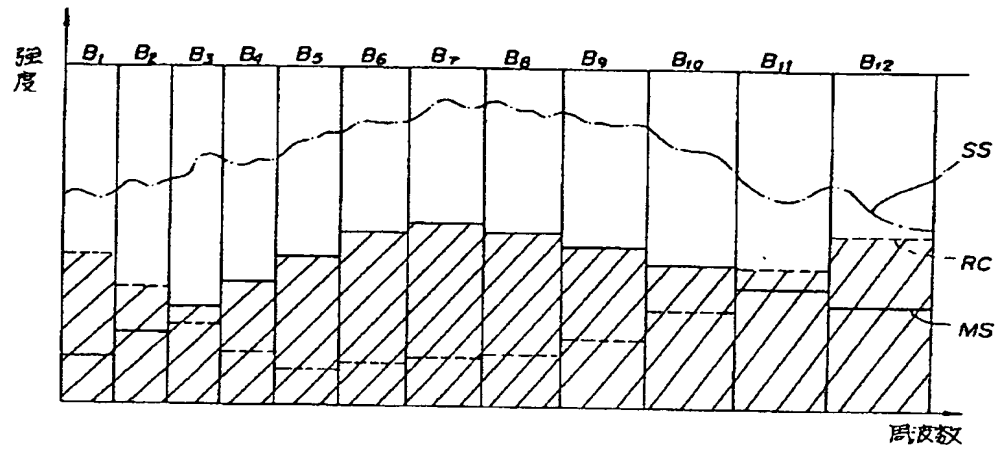


第 5 図



フィルタ回路

第 4 図



合成  
第 6 図